Producto Interno y Ortogonalidad

Departamento de Matemáticas, CSI/ITESM

15 de octubre de 2009

Índice

8.1. Contexto

Hemos mencionado que el problema fundamental del álgebra lineal consiste en resolver un sistema de ecuaciones lineales $\mathbf{A}\,\mathbf{x} = \mathbf{b}$ y hemos introducido varios conceptos sofisticados para manejar la solución y el análisis. El concepto de espacio generado resuelve el problema de la consistencia: El sistema es consistente si y sólo si $\mathbf{b} \in C(\mathbf{A})$ (el espacio generado por las columnas de $\mathbf{A}$). Otro concepto importante es el concepto de dependencia lineal y resuelve el problema de la unicidad: Si el sistema $\mathbf{A}\,\mathbf{x} = \mathbf{b}$ es consistente, entonces $\mathbf{A}\,\mathbf{x} = \mathbf{b}$ tiene infinitas soluciones si y sólo si las columnas de $\mathbf{A}$ forman un conjunto linealmente dependiente. Los conceptos de dimensión y de base de un espacio lineal permiten tener una idea qué tan grande es el conjunto solución y cómo generarlo: Si el sistema $\mathbf{A}\,\mathbf{x} = \mathbf{b}$ es consistente, entonces la fórmula de todas las soluciones es $\mathbf{y} = \mathbf{y}_o + \text{ns}(\mathbf{A})$ donde $\mathbf{y}_o$ es una solución particular y $\text{ns}(\mathbf{A})$ es el espacio nulo de $\mathbf{A}$, es decir, el conjunto de todos los vectores que satisfacen $\mathbf{A}\,\mathbf{x} = \mathbf{0}$. De hecho, si $\mathscr{B} = \{\mathbf{x}_1, \ldots, \mathbf{x}_r\}$ es una base para $\text{ns}(\mathbf{A})$, entonces la solución general será

$$\mathbf{y} = \mathbf{y}_o + \sum_{i=1}^{r} c_i \cdot \mathbf{x}_i$$

y esta será la solución general en su forma más reducida posible. El caso consistente ha dado origen a toda la teoría que hemos visto hasta ahora. Pasemos ahora a la inconsistencia. De alguna forma debemos cambiar la pregunta por que el tema se agotó: Si es $\mathbf{A}\,\mathbf{x} = \mathbf{b}$ inconsistente, entonces no hay solución. Una forma adecuada de reformular la pregunta es: No habiendo solución para $\mathbf{A}\,\mathbf{x} = \mathbf{b}$, ¿qué es lo más cerca posible que podemos

estar de una *solución*? Esto involucra dos conceptos hermanados: distancia y perpendicularidad. Y ellos tienen como origen el concepto de producto interno o producto punto que es el tema de esta lectura.

8.2. Introducción

En esta lectura veremos cómo definir un producto punto o producto interno en espacios de matrices (pues tenga presente que nos interesa el paso de $\mathbf{A}\,\mathbf{x} = \mathbf{b}$ a $\mathbf{A}\,\mathbf{X} = \mathbf{B}$ también en el caso inconsistente). Veremos que aunque en un principio la definición del producto punto apartir del uso de la traza sea *extraña*, esta definición es de lo más cómoda porque hace coincidir el producto punto entre matrices con el producto punto tradicional entre vectores cuando las matrices son vectorizadas. La estructura clásica de la construcción del Algebra Lineal referente al producto interno es: definir un producto interno, probar ciertas propiedades básicas, estas propiedades incluyen en concepto de ortogonalidad, y entonces mostrar que se puede definir una norma a partir del producto interno; para llegar la definir una distancia con la norma se requiere probar la desigualdad de Cauchy-Schwarz. Cuando se dice que se tiene una distancia en el espacio lineal se tiene el concepto de cercanía y el de error. Las propiedades importantes que deben ser deducidas para que la distancia definida sea cabalmente una distancia son de identidad (que si la distancia entre dos objetos es cero, entonces los objetos son idénticos), de simetría (que la distancia medida de un objeto a otro es independiente de cual de los dos sea el punto de referencia), la desigualdad del triángulo (la distancia entre dos puntos no excede la suma de las distancias de uno de ellos a otro intermedio mas la distancia de ese intermedio al segundo punto; y habrá igualdad cuando el punto intermedio esté en el segmento que une ambos puntos). El proceso de Gram-Schmidt será la clave para encontrar una base para $C(\mathbf{A})$ desde donde será fácil encontrar el punto más próximo de $C(\mathbf{A})$ a $\mathbf{b}$ y allí a la *mejor solución posible* a $\mathbf{A}\,\mathbf{x} = \mathbf{b}$.

8.3. Un producto punto mediante la traza

El producto punto entre matrices se define apartir del producto, la transpuesta y la traza entre y de matrices. En apariencia complejo, veremos su conveniencia y facilidad posteriormente.

Definición

Sean $\mathbf{A}$ y $\mathbf{B}$ dos matrices $m \times n$, el *producto interno* o *producto punto* de $\mathbf{A}$ y $\mathbf{B}$ se define como:

$$\mathbf{A} \bullet \mathbf{B} = \operatorname{tr}\left(\mathbf{A}' \cdot \mathbf{B}\right) \tag{1}$$

Ejemplifiquemos el cálculo de productos punto.

Ejemplo 1

Determine $\mathbf{A} \bullet \mathbf{B}$ si

$$\mathbf{A} = \begin{bmatrix} 3 & -1 \\ 2 & 0 \\ 5 & -2 \end{bmatrix}, \ \mathbf{B} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \\ 4 & -3 \end{bmatrix}$$

Solución

$$\mathbf{A}' \cdot \mathbf{B} = \begin{bmatrix} 3 & 2 & 5 \\ -1 & 0 & -2 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & 1 \\ 4 & -3 \end{bmatrix} = \begin{bmatrix} 23 & -16 \\ -9 & 7 \end{bmatrix}$$

Por tanto:

$$\mathbf{A} \bullet \mathbf{B} = \operatorname{tr}\left(\begin{bmatrix} 23 & -16 \\ -9 & 7 \end{bmatrix}\right) = 23 + 7 = 30 \ \blacksquare$$

Toca el turno al alumno para ejercitar el concepto del producto punto entre matrices de la misma dimensión.

Ejercicio 1

Para las matrices siguientes calcule $\mathbf{A} \bullet \mathbf{B}$.

1. $\mathbf{A} = \begin{bmatrix} 2 & -1 \\ 3 & 1 \end{bmatrix}$ y $\mathbf{B} = \begin{bmatrix} 1 & 0 \\ -1 & 0 \end{bmatrix}$

2. $\mathbf{A} = \begin{bmatrix} 2 & -1 & 3 \\ 1 & 0 & 1 \end{bmatrix}$ y $\mathbf{B} = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 2 & 1 \end{bmatrix}$

3. $\mathbf{A} = \begin{bmatrix} 2 & -1 \\ 3 & 1 \\ 0 & 1 \end{bmatrix}$ y $\mathbf{B} = \begin{bmatrix} 1 & 0 \\ -1 & 0 \\ 2 & 1 \end{bmatrix}$

8.4. Propiedades del producto punto

El camino tradicional para ver que un producto es efectivamente candidato a producto interno con escalares reales es probar que es: conmutativo, no negativo, y *bilineal*:

Teorema 8.1

Sean $\mathbf{A}$, $\mathbf{B}$ y $\mathbf{C}$ matrices $m \times n$ y sea k un escalar. Entonces el producto punto cumple las siguientes propiedades:

1. Simetría: $\mathbf{A} \bullet \mathbf{B} = \mathbf{B} \bullet \mathbf{A}$
2. No negatividad: $\mathbf{A} \bullet \mathbf{A} \geq 0$, habiendo igualdad *ssi* $\mathbf{A} = \mathbf{0}$
3. Proporcionalidad: $(k\mathbf{A}) \bullet \mathbf{B} = k\,(\mathbf{A} \bullet \mathbf{B})$
4. Distributividad: $(\mathbf{A} + \mathbf{B}) \bullet \mathbf{C} = \mathbf{A} \bullet \mathbf{C} + \mathbf{B} \bullet \mathbf{C}$
5. $\mathbf{0} \bullet \mathbf{A} = 0$

Demostración

1. Directo de la definición:

$$\begin{aligned} \mathbf{A} \bullet \mathbf{B} &= \operatorname{tr}(\mathbf{A}'\mathbf{B}) \\ &= \operatorname{tr}\left((\mathbf{A}'\mathbf{B})'\right) \\ &= \operatorname{tr}(\mathbf{B}'\mathbf{A}) \\ &= \mathbf{B} \bullet \mathbf{A} \end{aligned}$$

2. De la sección anterior:

$$\begin{aligned} \mathbf{A} \bullet \mathbf{A} &= \operatorname{tr}(\mathbf{A}'\mathbf{A}) \\ &= \textstyle\sum_{i=1}^m \sum_{j=1}^n a_{ij}^2 \geq 0 \end{aligned}$$

Además, $\mathbf{A} \bullet \mathbf{A} = 0$ ssi

$$\sum_{i=1}^m \sum_{j=1}^n a_{ij}^2 = 0$$

ssi $a_{ij} = 0$ para todo $i = 1 \ldots m$ y $j = 1 \ldots n$ ssi $\mathbf{A} = \mathbf{0}$ ■

No sea impaciente; Ud. participará en el *festín* de mini-demostraciones:

Ejercicio 2

Demuestre el punto 3 del teorema anterior:

Sean $\mathbf{A}$ y $\mathbf{B}$ matrices $m \times n$ y sea k un escalar. Pruebe que:

$$(k\mathbf{A}) \bullet \mathbf{B} = k\,(\mathbf{A} \bullet \mathbf{B})$$

Sugerencia

Utilice la definición de $\bullet$, la propiedad de la transpuesta, y la propiedad 1 del Lema 5.1.

Ejercicio 3

Demuestre el punto 4 del teorema anterior:
Sean **A**, **B** y **C** matrices $m \times n$. Pruebe que:

$$(\mathbf{A} + \mathbf{B}) \bullet \mathbf{C} = \mathbf{A} \bullet \mathbf{C} + \mathbf{B} \bullet \mathbf{C}$$

Sugerencia
Utilice la definición de $\bullet$, la propiedad de la transpuesta, y la propiedad 2 del Lema 5.1.

Ejercicio 4

Demuestre el punto 5 del teorema anterior:
Sea **A** una matriz $m \times n$. Pruebe que:

$$\mathbf{0} \bullet \mathbf{A} = 0$$

8.5. La norma de una matriz

Habiendo probado las propiedades mencionadas del producto interno nuevo, el siguiente paso es definir la norma de una matriz. En el espacio de matrices hay un buen número de normas igualmente útiles, la norma que se define aquí es la *norma-2 de una matriz*.

Definición
La *norma* de la matriz **A** $m \times n$ se define como:

$$\|\mathbf{A}\|_2 = \|\mathbf{A}\| = \sqrt{\mathbf{A} \bullet \mathbf{A}} \qquad (2)$$

Observe que por el punto 2 del teorema anterior, $\mathbf{A} \bullet \mathbf{A} \geq 0$; así que la raíz cuadrada entregará un real mayor o igual a cero. Hagamos un ejemplo del cálculo de la norma de una matriz.

Ejemplo 2
Determine la norma de la matriz:

$$\mathbf{A} = \begin{bmatrix} 3 & -1 \\ 2 & 0 \\ 5 & -2 \end{bmatrix}$$

Solución Como

$$\mathbf{A}'\mathbf{A} = \begin{bmatrix} 3 & 2 & 5 \\ -1 & 0 & -2 \end{bmatrix} \begin{bmatrix} 3 & -1 \\ 2 & 0 \\ 5 & -2 \end{bmatrix} = \begin{bmatrix} 38 & -13 \\ -13 & 5 \end{bmatrix}$$

Así

$$\mathbf{A} \bullet \mathbf{A} = \mathrm{tr} \begin{bmatrix} 38 & -13 \\ -13 & 5 \end{bmatrix} = 43$$

y por tanto, $\|\mathbf{A}\| = \sqrt{43}$ ■
Es su turno de practicar el cálculo de la norma de una matriz.

Ejercicio 5

Para las matrices siguientes calcule su norma.

1. $\mathbf{A} = \begin{bmatrix} 2 & -1 \\ 3 & 1 \end{bmatrix}$

2. $\mathbf{A} = \begin{bmatrix} 2 & -1 & 3 \\ 1 & 0 & 1 \end{bmatrix}$

3. $\mathbf{A} = \begin{bmatrix} 2 & -1 \\ 3 & 1 \\ 0 & 1 \end{bmatrix}$

Un hecho importante sobre la norma de una matriz es:

$$\|\mathbf{A}\| = \operatorname{tr}\left(\mathbf{A}'\mathbf{A}\right) = \left(\sum_{i=1}^{m}\sum_{j=1}^{n}(a_{ij})^2\right)^{1/2}$$

Lo cual hace coincidir la definición tradicional de norma de un vector con la norma de una matriz cuando la matriz es utilizada para formar un solo vector formado por los renglones de ella. **Algo más importante de observar es que el producto punto entre matrices usando la traza equivale al producto punto tradicional aplicado a los vectores obtenidos de vectorizar las matrices.**

Ejemplo 3

Verifiquemos la afirmación en el caso de matrices 2×2.

Comprobación

Digamos que

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \mathbf{B} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$$

Así

$$\mathbf{A}'\mathbf{B} = \left[\begin{array}{c|c} a_{11}b_{11} + a_{21}b_{21} & ____ \\ \hline ____ & a_{12}b_{12} + a_{22}b_{22} \end{array}\right]$$

De donde

$$\mathbf{A} \bullet \mathbf{B} = (a_{11}b_{11} + a_{21}b_{21}) + (a_{12}b_{12} + a_{22}b_{22})$$

8.6. Propiedades de la norma de matrices

Habiendo definido la norma de una matriz, lo siguiente es probar algunas propiedades básicas.

Teorema 8.2

Sea $\mathbf{A}$ una matriz $m \times n$ y k un escalar:

1. $\|\mathbf{A}\| \geq 0$, además $\|\mathbf{A}\| = 0$ *ssi* $\mathbf{A} = \mathbf{0}$
2. $\|-\mathbf{A}\| = \|\mathbf{A}\|$
3. $\|k\mathbf{A}\| = |k|\,\|\mathbf{A}\|$

Demostración

1. De lo visto

$$\|\mathbf{A}\| = \sqrt{\mathbf{A} \bullet \mathbf{A}} = \left(\sum_{i=1}^{m}\sum_{j=1}^{n}(a_{ij})^2\right)^{1/2} \geq 0$$

Y también de lo mismo se deduce que $\|\mathbf{A}\| = 0$ ssi $\mathbf{A} = \mathbf{0}$.
3. Por las propiedades conocidas:

$$\|k\mathbf{A}\| = \sqrt{(k\mathbf{A}) \bullet (k\mathbf{A})} = \sqrt{k^2 \mathbf{A} \bullet \mathbf{A}} = |k| \sqrt{\mathbf{A} \bullet \mathbf{A}} = |k| \; \|\mathbf{A}\|$$

El segundo punto es el más sencillo y le queda hacerlo como ejercicio ■

Ejercicio 6

Demuestre el punto 2. del teorema anterior:
Sea $\mathbf{A}$ una matriz $m \times n$. Pruebe que:

$$\| - \mathbf{A}\| = \|\mathbf{A}\|$$

Sugerencia
Use directamente la definición de la magnitud de una matriz.

8.7. Desigualdad de Schwarz

En la línea tradicional, el siguiente paso para ver que nuestro producto punto llega a definir una distancia a través de la definición de la norma, es probar que nuestro producto punto y la norma cumplen una relación de desigualdad llamada la desigualdad de Cauchy-Schwarz.

Teorema 8.3

Sean $\mathbf{A}$ y $\mathbf{B}$ matrices $m \times n$:

$$|\mathbf{A} \bullet \mathbf{B}| \leq \|\mathbf{A}\| \cdot \|\mathbf{B}\| \tag{3}$$

Además, la igualdad se tiene si y sólo si una matriz es un múltiplo de la otra.

Demostración
En esta demostración es importante que vigile el rol de los factores en los productos: en los productos $\bullet$ los factores en ambos lados deben ser matrices, mientras en los productos $\cdot$ el factor a la izquierda es escalar y el factor a la derecha debe ser una matriz. Para cualesquiera escalares x y y:

$$\begin{array}{rcl} 0 & \leq & (x \cdot \mathbf{A} - y \cdot \mathbf{B}) \bullet (x \cdot \mathbf{A} - y \cdot \mathbf{B}) \\ & = & x^2 \|\mathbf{A}\|^2 - 2xy\,(\mathbf{A} \bullet \mathbf{B}) + y^2 \|\mathbf{B}\|^2 \end{array}$$

Suponiendo $\mathbf{B} \neq \mathbf{0}$, tomamos $x = \|\mathbf{B}\|$ y $y = (\mathbf{A} \bullet \mathbf{B})/\|\mathbf{B}\|$ en la fórmula anterior y simplificando muchos términos obtenemos:

$$0 \leq \|\mathbf{B}\|^2 \, \|\mathbf{A}\|^2 - (\mathbf{A} \bullet \mathbf{B})^2$$

De donde se obtiene:

$$(\mathbf{A} \bullet \mathbf{B})^2 \leq \|\mathbf{B}\|^2 \, \|\mathbf{A}\|^2$$

y tomando raíz cuadrada se obtiene la conclusión. Recuerde que un real k, $\sqrt{k^2} = |k|$ y no que $\sqrt{k^2} = k$ pues para los valores de k negativos *el signo se pierde al elevar al cuadrado.* Además, si $\mathbf{A} = k \cdot \mathbf{B}$ se tiene que $\|\mathbf{A}\| = |k| \, \|\mathbf{B}\|$ y así

$$|\mathbf{A} \bullet \mathbf{B}| = |k| \, (\mathbf{B} \bullet \mathbf{B}) = |k| \, \|\mathbf{B}\|^2 = \|\mathbf{A}\| \, \|\mathbf{B}\|$$

probado que la desigualdad se cumple con igualdad en este caso. Por otro lado, si $|\mathbf{A} \bullet \mathbf{B}| = \|\mathbf{A}\| \, \|\mathbf{B}\|$ definimos

$$\mathbf{C} = \|\mathbf{B}\|^2 \cdot \mathbf{A} - (\mathbf{A} \bullet \mathbf{B}) \cdot \mathbf{B}$$

se puede confirmar que desarrollando el producto término a término: $\mathbf{C} \bullet \mathbf{C} = 0$, implicando que $\mathbf{C} = \mathbf{0}$, de donde $\|\mathbf{B}\|^2 \cdot \mathbf{A} = (\mathbf{A} \bullet \mathbf{B}) \cdot \mathbf{B}$ si suponemos $\mathbf{B} \neq \mathbf{0}$ deducimos que

$$\mathbf{A} = \frac{(\mathbf{A} \bullet \mathbf{B})}{\|\mathbf{B}\|^2} \cdot \mathbf{B}$$

lo que dice que $\mathbf{A}$ es un múltiplo de $\mathbf{B}$. Si $\mathbf{B} = \mathbf{0}$, claramente $\mathbf{B} = 0 \cdot \mathbf{A}$ y la desigualdad se cumple con igualdad ■

Ejercicio 7

Sean $\mathbf{A}$ y $\mathbf{B}$ dos matrices $m \times n$. Demuestre que

$$\|\mathbf{A} + \mathbf{B}\| \leq \|\mathbf{A}\| + \|\mathbf{B}\|$$

y que existe igualdad sólo cuando una matriz es un múltiplo escalar de la otra.

Sugerencia

Utilice

$$\|\mathbf{A} + \mathbf{B}\|^2 = (\mathbf{A} + \mathbf{B}) \bullet (\mathbf{A} + \mathbf{B})$$

Desarrolle tal producto punto. Después utilice la desigualdad de Schwarz. Finalmente, tome raíz cuadrada. Para la igualdad se requiere igualdad en la desigualdad de Schwarz.

Lema 8.4

Sean $\mathbf{X}$ y $\mathbf{Y}$ dos matrices $m \times n$. Entonces $\mathbf{X}$ y $\mathbf{Y}$ son ortogonales, es decir, cumplen $\mathbf{X} \bullet \mathbf{Y} = 0$ si y sólo si

$$\|\mathbf{X} \pm \mathbf{Y}\|^2 = \|\mathbf{X}\|^2 + \|\mathbf{Y}\|^2$$

Demostración

Tenemos en general que

$$\begin{aligned} \|\mathbf{X} \pm \mathbf{Y}\|^2 &= (\mathbf{X} \pm \mathbf{Y}) \bullet (\mathbf{X} \pm \mathbf{Y}) \\ &= \mathbf{X} \bullet \mathbf{X} \pm 2\,\mathbf{X} \bullet \mathbf{Y} + \mathbf{Y} \bullet \mathbf{Y} \\ &= \|\mathbf{X}\|^2 + \|\mathbf{Y}\|^2 \pm 2\,\mathbf{X} \bullet \mathbf{Y} \end{aligned}$$

Por tanto, si $\mathbf{X} \bullet \mathbf{Y} = 0$ entonces de la fórmula anterior se deduce la igualdad que queremos probar. Por otro lado, si la fórmula se cumple nuestra fórmula indica que $\mathbf{X} \bullet \mathbf{Y} = 0$. Es decir $\mathbf{X}$ y $\mathbf{Y}$ son perpendiculares. ■

8.8. La distancia entre dos matrices

Habiendo probado la desigualdad de Cauchy-Schwarz se puede construir una función distancia entre el espacio de matrices:

Definición

Sean $\mathbf{A}$ y $\mathbf{B}$ dos matrices $m \times n$ la *distancia entre matrices* $\mathbf{A}$ y $\mathbf{B}$ se define como:

$$\boxed{\delta\left(\mathbf{A}, \mathbf{B}\right) = \|\mathbf{A} - \mathbf{B}\|} \qquad (4)$$

Realicemos un ejemplo simple para ilustrar este concepto.

Determine la distancia entre las matrices:

$$\mathbf{A} = \begin{bmatrix} 3 & -1 \\ 2 & 0 \\ 5 & -2 \end{bmatrix} \; \mathbf{B} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \\ 1 & -2 \end{bmatrix}$$

Solución Si

$$\mathbf{C}=\mathbf{A}-\mathbf{B}=\begin{bmatrix} 2 & 0 \\ 2 & -1 \\ 4 & 0 \end{bmatrix}$$

$$\mathbf{C}'\mathbf{C}=\begin{bmatrix} 2 & 2 & 4 \\ 0 & -1 & 0 \end{bmatrix}\begin{bmatrix} 2 & 0 \\ 2 & -1 \\ 4 & 0 \end{bmatrix}=\begin{bmatrix} 24 & -2 \\ -2 & 1 \end{bmatrix}$$

Así

$$\delta(\mathbf{A},\mathbf{B})=\|\mathbf{C}\|=\sqrt{\mathbf{C}\bullet\mathbf{C}}=\sqrt{\operatorname{tr}(\mathbf{C}'\mathbf{C})}=\sqrt{24+1}=5_\diamond$$

Ejercicio 8

Determine la distancia entre las matrices:

1. $\mathbf{A}=\begin{bmatrix} 3 & -1 \\ 2 & 1 \\ 5 & 2 \end{bmatrix}$ y $\mathbf{B}=\begin{bmatrix} 1 & 1 \\ -1 & 1 \\ 1 & 2 \end{bmatrix}$
2. $\mathbf{A}=\begin{bmatrix} 3 & -1 \\ 5 & 2 \end{bmatrix}$ y $\mathbf{B}=\begin{bmatrix} -1 & 1 \\ 1 & 2 \end{bmatrix}$

Ejercicio 9

Sean $\mathbf{A}$ y $\mathbf{B}$ dos matrices de iguales dimensiones. Pruebe que:

$$\delta(\mathbf{A},\mathbf{B})=\delta(\mathbf{B},\mathbf{A})$$

Sugerencia
Use directamente la definición de distancia entre matrices. Utilice también la propiedad conmutativa del producto punto.

Ejercicio 10

Sean $\mathbf{A}$ y $\mathbf{B}$ dos matrices de iguales dimensiones. Pruebe que:

$$\delta(\mathbf{A},\mathbf{B})\geq 0$$

Además, pruebe que hay igualdad si y sólo si $\mathbf{A}=\mathbf{B}$.
Sugerencia
Utilice la definición de distancia entre matrices y la propiedad 1 del teorema 6.2 sobre la norma para matrices.

Ejercicio 11

Sean $\mathbf{A}$, $\mathbf{B}$ y $\mathbf{C}$ dos matrices de iguales dimensiones. Pruebe que:

$$\delta(\mathbf{A},\mathbf{B})\leq\delta(\mathbf{A},\mathbf{C})+\delta(\mathbf{C},\mathbf{B})$$

Vea que existe igualdad si sólo si existe un escalar t $(0\leq t\leq 1)$ tal que $\mathbf{C}=t\,\mathbf{A}+(1-t)\,\mathbf{B}$. Esta desigualdad se llama *Desigualdad del triángulo*.
Sugerencia
Utilice como válido el resultado del ejercicio 7 con $\mathbf{A}$ como $\mathbf{A}-\mathbf{C}$ y con $\mathbf{B}$ como $\mathbf{C}-\mathbf{B}$.

Ejercicio 12

Sean $\mathbf{A}$, $\mathbf{B}$ y $\mathbf{C}$ dos matrices de iguales dimensiones. Pruebe que:

$$\delta(\mathbf{A},\mathbf{B}) = \delta(\mathbf{A}+\mathbf{C},\mathbf{B}+\mathbf{C})$$

Esta propiedad indica que la distancia entre matrices se *preserva ante traslaciones*.

Sugerencia

Directamente de la definición de distancia.

8.9. Ángulo entre matrices

Definición

Sean $\mathbf{A}$ y $\mathbf{B}$ dos matrices $m \times n$, no nulas, el *ángulo entre matrices* $\mathbf{A}$ y $\mathbf{B}$, θ, se define como

$$\boxed{\cos(\theta) = \frac{\mathbf{A}\bullet\mathbf{B}}{\|\mathbf{A}\|\,\|\mathbf{B}\|}} \tag{5}$$

Aquí θ debe cumplir: $0 \leq \theta \leq \pi$.

Ejemplo 4

Determine el ángulo en radianes entre las matrices

$$\mathbf{A} = \begin{bmatrix} 3 & -1 \\ 2 & 0 \\ 5 & -2 \end{bmatrix},\ \mathbf{B} = \begin{bmatrix} -4 & -1 \\ 0 & 0 \\ 1 & -2 \end{bmatrix}$$

Solución Como

$$\mathbf{A}\bullet\mathbf{B} = -2,\ \|\mathbf{A}\| = \sqrt{43},\ \|\mathbf{B}\| = \sqrt{22}$$

Entonces

$$\cos(\theta) = \frac{-2}{\sqrt{43}\sqrt{22}} \approx -0.065,\ \text{ por tanto } \theta \approx 1.63\,\text{radianes}_\diamond$$

Ejercicio 13

Determine el ángulo en radianes entre las matrices

$$\mathbf{A} = \begin{bmatrix} 3 & -1 \\ 5 & -2 \end{bmatrix},\ \mathbf{B} = \begin{bmatrix} -4 & -1 \\ 1 & -2 \end{bmatrix}$$

Ejercicio 14

Determine el valor de x para que el ángulo en radianes entre las matrices

$$\mathbf{A} = \begin{bmatrix} 3 & -1 \\ 5 & -2 \end{bmatrix},\ \mathbf{B} = \begin{bmatrix} -4 & -1 \\ 1 & x \end{bmatrix}$$

sea de 0.1.

Sugerencia

Entréle sin miedo: $\mathbf{A}\bullet\mathbf{B}$ es una expresión lineal en x. $|\mathbf{B}|$ es una raíz cuadrada de un polinomio cuadrático en x. Cuando se iguala:

$$\cos(0.1\text{rad}) = \frac{\mathbf{A}\bullet\mathbf{B}}{\|\mathbf{A}\|\,\|\mathbf{B}\|}$$

después de elevar al cuadrado y multiplicar por denominadores queda una ecuación cuadrática en x.

8.10. Ortogonalidad y espacios generados

Sean $\mathbf{A}$ y $\mathbf{B}$ dos matrices $m \times n$, se dice que son *matrices ortogonales* si se cumple

$$\mathbf{A} \bullet \mathbf{B} = 0$$

Un conjunto $\{\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_k\}$ de matrices $m \times n$ se dice *conjunto ortogonal* si $\mathbf{A}_i \bullet \mathbf{A}_j = 0$ para $i \neq j$ El conjunto anterior se dice *ortonormal* si es ortogonal y además cada matriz tiene norma 1. Una vez que se tiene un conjunto ortogonal de elementos no cero

$$\{\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_k\}$$

es posible convertirlo a uno ortonormal

$$\{\mathbf{B}_1, \mathbf{B}_2, \ldots, \mathbf{B}_k\}$$

definiendo:

$$\mathbf{B}_i = \frac{1}{\|\mathbf{A}_i\|}\mathbf{A}_i$$

Ejercicio 15

Determine el valor de x para que las matrices sean ortogonales.

$$\mathbf{A} = \begin{bmatrix} 3 & -1 \\ 5 & -2 \end{bmatrix}, \ \mathbf{B} = \begin{bmatrix} -4 & -1 \\ 1 & x \end{bmatrix}$$

Sugerencia
Al hacer $\mathbf{A} \bullet \mathbf{B} = 0$ queda una ecuación lineal en x.

Ejercicio 16

Determine qué condición deben cumplir las variables para que $\mathbf{A} \perp \mathbf{B}$.

$$\mathbf{A} = \begin{bmatrix} 3 & -1 \\ 5 & -2 \end{bmatrix}, \ \mathbf{B} = \begin{bmatrix} x & y \\ z & w \end{bmatrix}$$

Sugerencia
Al hacer $\mathbf{A} \bullet \mathbf{B} = 0$ queda una ecuación lineal.

Ejercicio 17

Sea $\mathbf{A}$ una matriz $m \times n$ muestre que el conjunto

$$\mathbf{A}^{\perp} = \{\mathbf{X} \text{ matriz } m \times n | \mathbf{A} \perp \mathbf{X}\}$$

es un espacio lineal.
Sugerencia
Al hacer $\mathbf{A} \bullet \mathbf{X} = 0$ queda una ecuación homogénea. Las soluciones a los sistemas homogéneos son espacios lineales.

Ejercicio 18

Determine el espacio perpendicular a la matriz:

$$\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

Escíbalo como un espacio generado.
Sugerencia
Tome una matriz $\mathbf{X}$ 2×2 con incógnitas. Al hacer $\mathbf{A} \bullet \mathbf{X} = 0$ queda una ecuación homogénea con 4 incógnitas; una de ellas es fija y las otras son libres. Despeje la fija y póngala en función de las libres. En lugar de la notación de vector en la solución general, prefiera la notación de matriz.

Ejercicio 19

Sea V un subespacio lineal de matrices $m \times n$ muestre que el conjunto

$$V^{\perp} = \{\mathbf{X} \text{ matriz } m \times n | \mathbf{Y} \perp \mathbf{X} \text{ para toda } \mathbf{Y} \in V\}$$

es un espacio lineal.
Sugerencia
Muestre que

- no es vacío: la matriz de ceros está (compruébelo),
- es cerrado bajo la suma (compruébelo), y
- es cerrado bajo el producto por escalares (compruébelo).

Ejercicio 20

Determine el espacio perpendicular al espacio generado por las matrices matrices:

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \text{ y } \mathbf{B} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

Sugerencia
Tome una matriz $\mathbf{X}$ 2×2 de incógnitas:

- $\mathbf{A} \bullet \mathbf{X} = 0$ da una ecuación en las incógnitas
- $\mathbf{B} \bullet \mathbf{X} = 0$ da otra ecuación en las incógnitas

Resuelva el sistema homogéneo, encuentre la solución general y déle la notación matricial.

Veamos ahora un par de resultados teóricos:

Lema 8.5

Un conjunto ortogonal de matrices no nulas es linealmente independiente

Demostración
Si $\{\mathbf{A}_1 \mathbf{A}_2, \ldots, \mathbf{A}_k\}$ es un conjunto ortogonal de matrices no nulas y si se cumple:

$$c_1 \mathbf{A}_1 + c_2 \mathbf{A}_2 + \cdots + c_k \mathbf{A}_k = \mathbf{0}$$

veamos que los coeficientes c_i son cero. Haciendo el producto punto con $\mathbf{A}_i$:

$$\mathbf{A}_i \bullet (c_1\mathbf{A}_1 + c_2\mathbf{A}_2 + \cdots + c_k\mathbf{A}_k) = \mathbf{A}_i \bullet \mathbf{0} = 0$$

Desarrollando el primer miembro:

$$\mathbf{A}_i \bullet (c_1\mathbf{A}_1) + \mathbf{A}_i \bullet (c_2\mathbf{A}_2) + \cdots + \mathbf{A}_i \bullet (c_k\mathbf{A}_k) = 0$$

Como el conjunto es ortogonal entonces $\mathbf{A}_i \bullet \mathbf{A}_j = 0$ para $i \neq j$, la expresión anterior queda:

$$c_i\mathbf{A}_i \bullet \mathbf{A}_i = 0$$

Puesto que ninguna matriz $\mathbf{A}_i$ es nula, $\mathbf{A}_i \bullet \mathbf{A}_i \neq 0$,la expresión anterior queda:

$$c_i = 0$$

Así, cada coeficiente en la combinación lineal es cero. Por tanto, el conjunto de matrices es linealmente independiente.■

Cuando se quiere verificar que una matriz es ortogonal al generado por un conjunto de matrices enfrentámos un dilema singular, cómo hacer pues en el generado resulta en conjuntos infinitos? El siguiente resultado es importante porque reduce una verificación infinita a una verificación sobre el conjunto generador:

Lema 8.6

Sea $\mathbf{A}$ una matriz y $V = \text{Gen}\{\mathbf{B}_1, \ldots, \mathbf{B}_k\}$ en $\mathbf{M}_{m\times n}$. Entonces: $\mathbf{A}$ es ortogonal a todo V si y sólo si $\mathbf{A}$ es ortogonal a cada $\mathbf{B}_i$.

Demostración

Suficiencia: Si suponemos que $\mathbf{A}$ es ortogonal a todo V y sabiendo que cada $\mathbf{B}_i$ está en V, entonces es inmediato que $\mathbf{A}$ es ortogonal a $\mathbf{B}_i$.

Necesidad: Supongamos ahora que $\mathbf{A}$ es ortogonal a todo $\mathbf{B}_i$, es decir que $A \bullet \mathbf{B}_i = 0$ para todo $i = 1, \ldots, k$. Veamos que entonces es ortogonal a cualquier $\mathbf{X}$ de V. Tomemos un $\mathbf{X}$ cualquiera de V. Como V está generado por los $\mathbf{B}_i$, entonces deben existir escalares c_i tales que

$$\mathbf{X} = \sum_{i=1}^{k} c_k \cdot \mathbf{B}_i$$

Por tanto,

$$\begin{array}{rcl} \mathbf{A} \bullet \mathbf{X} & = & \mathbf{A} \bullet (\sum_{i=1}^{k} c_k \cdot \mathbf{B}_i) \\ & = & \sum_{i=1}^{k} \mathbf{A} \bullet (c_k \cdot \mathbf{B}_i) \\ & = & \sum_{i=1}^{k} c_i(\mathbf{A} \bullet \mathbf{B}_i) \\ & = & \sum_{i=1}^{k} c_i(0) \\ & = & 0 \end{array}$$

Por tanto, $\mathbf{A}$ es ortogonal a $\mathbf{X}$ ■

8.11. Bases Ortonormales

El siguiente resultado es importante porque permite cambiar la base de un espacio generado por otra que es ortogonal. La misma demostración del resultado da el algoritmo de conversión.

Teorema 8.7

Sea un conjunto $\{\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_k\}$ de matrices $m \times n$ linealmente independiente. Entonces existen escalares únicos $x_{i,j} (i < j)$ tales que:

$$\begin{array}{rcl} \mathbf{B}_1 & = & \mathbf{A}_1 \\ \mathbf{B}_2 & = & \mathbf{A}_2 - x_{1,2}\mathbf{B}_1 \\ & \vdots & \\ \mathbf{B}_j & = & \mathbf{A}_j - x_{1,j}\mathbf{B}_1 - x_{2,j}\mathbf{B}_2 - \cdots - x_{j-1,j}\mathbf{B}_{j-1} \\ & \vdots & \\ \mathbf{B}_k & = & \mathbf{A}_k - x_{1,k}\mathbf{B}_1 - x_{2,k}\mathbf{B}_2 - \cdots - x_{k-1,k}\mathbf{B}_{k-1} \end{array}$$

forman un conjunto ortogonal.

Demostración

Definimos los valores escalares $x_{i,j}$ como precisamente se requieren:

$$x_{i,j} = \frac{\mathbf{A}_j \bullet \mathbf{B}_i}{\mathbf{B}_i \bullet \mathbf{B}_i} \qquad (6)$$

Ud. puede recordarlos si ubica que el primer subíndice de x_{ij} se relaciona a las matrices por construir $\mathbf{B}_i$ y que ésta es la que aparece en el denominador con el producto punto con si misma. Un punto importante a observar es que los $\{\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_j\}$ son combinaciones lineales de los $\{\mathbf{B}_1, \mathbf{B}_2, \ldots, \mathbf{B}_j\}$ y por consiguiente el conjunto de los $\mathbf{B}s$ es linealmente independiente. Para verificar que el conjunto es ahora ortogonal hay que realizar los productos punto. Conviene hacer una demostración inductiva de la afirmación. La afirmación es que el conjunto $\{\mathbf{B}_1, \ldots, \mathbf{B}_k\}$ es ortogonal. Esta afirmación sería la afirmación última de la cadena de afirmaciones $\{\mathbf{B}_1, \ldots, \mathbf{B}_i\}$ para $i = 2, 3, \ldots, k$. La base inductiva para $i = 2$ se comprueba revisando que $\mathbf{B}_1$ y $\mathbf{B}_2$ son ortogonales:

$$\mathbf{B}_2 \bullet \mathbf{B}_1 = \mathbf{A}_2 \bullet \mathbf{B}_1 - x_{1,2}\mathbf{B}_1 \bullet \mathbf{B}_1 = \mathbf{A}_2 \bullet \mathbf{B}_1 - \frac{\mathbf{A}_2 \bullet \mathbf{B}_1}{\mathbf{B}_1 \bullet \mathbf{B}_1}\mathbf{B}_1 \bullet \mathbf{B}_1 = 0$$

Supongamos que $\{\mathbf{B}_1, \ldots, \mathbf{B}_i\}$ es un conjunto ortogonal para $1 \leq i \leq k-1$, veamos $\{\mathbf{B}_1, \ldots, \mathbf{B}_i, \mathbf{B}_{i+1}\}$ también lo será. Para ello basta probar que $\mathbf{B}_{i+1}$ es ortogonal a cada $\mathbf{B}_j$ para $1 \leq j \leq i$. Como

$$\mathbf{B}_{i+1} = \mathbf{A}_{i+1} - \sum_{l=1}^{i} x_{l,i+1}\mathbf{B}_l$$

Así

$$\begin{array}{rcl} \mathbf{B}_{i+1} \bullet \mathbf{B}_j & = & \left(\mathbf{A}_{i+1} - \sum_{l=1}^{i} x_{l,i+1}\mathbf{B}_l\right) \bullet \mathbf{B}_j \\ & = & \mathbf{A}_{i+1} \bullet \mathbf{B}_j - \sum_{l=1}^{i} x_{l,i+1}\mathbf{B}_l \bullet \mathbf{B}_j \\ & = & \mathbf{A}_{i+1} \bullet \mathbf{B}_j - x_{j,i+1}\mathbf{B}_j \bullet \mathbf{B}_j \\ & = & 0 \end{array}$$

Note que en el penúltimo paso fue requerida la base inductiva: que $\mathbf{B}_j$ era ortogonal a sus compañeros en $\{\mathbf{B}_1, \ldots, \mathbf{B}_i\}$.■

Ejemplo 5

Aplique el algoritmo anterior al conjunto formado por las matrices:

$$\mathbf{A}_1 = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}, \mathbf{A}_2 = \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix}, \mathbf{A}_3 = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}$$

Solución

$$\mathbf{B}_1 = \mathbf{A}_1$$

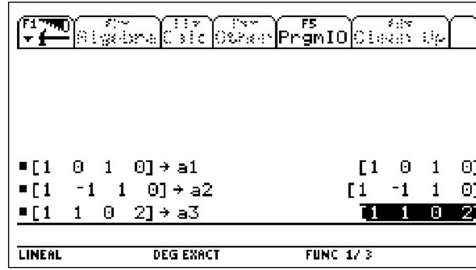


Figura 1: Captura de la matrices vectorizadas del ejemplo 7.

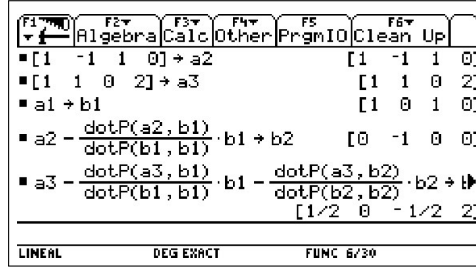


Figura 2: Obtención del conjunto ortogonal del ejemplo 7.

$$\mathbf{B}_1 \bullet \mathbf{B}_1 = 2 \text{ y } \mathbf{A}_2 \bullet \mathbf{B}_1 = 2; x_{12} = 1$$

$$\mathbf{B}_2 = \mathbf{A}_2 - x_{12}\mathbf{B}_1 = \begin{bmatrix} 0 & -1 \\ 0 & 0 \end{bmatrix}$$

Realizando los cálculos:

$$\mathbf{B}_2 \bullet \mathbf{B}_2 = 1; \mathbf{A}_3 \bullet \mathbf{B}_1 = 1; \mathbf{A}_3 \bullet \mathbf{B}_2 = -1;$$

Obtenemos:

$$x_{13} = 1/2, x_{23} = -1;$$

Y así:

$$\mathbf{B}_3 = \mathbf{A}_3 - x_{13}\mathbf{B}_1 - x_{23}\mathbf{B}_2 = \begin{bmatrix} 1/2 & 0 \\ -1/2 & 2 \end{bmatrix}_\diamond$$

Estos cálculos pueden hacerse en la calculadora TI, para ello habrá que vectorizar las matrices y seguir las fórmulas: en la figura 1 aparecen capturadas las matrices como vectores y en la figura 2 aparecen los cálculos del algoritmo de ortogonalización.

Ejemplo 6

Aplique el algoritmo anterior al conjunto formado por las matrices:

$$\mathbf{A}_1 = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}, \mathbf{A}_2 = \begin{bmatrix} 2 & 0 \\ 2 & 0 \end{bmatrix}, \mathbf{A}_3 = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}$$

Solución

$$\mathbf{B}_1 = \mathbf{A}_1$$

$$\mathbf{B}_1 \bullet \mathbf{B}_1 = 2 \text{ y } \mathbf{A}_2 \bullet \mathbf{B}_1 = 4; x_{12} = 2$$

$$\mathbf{B}_2 = \mathbf{A}_2 - x_{12}\mathbf{B}_1 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

este cálculo indica que vector $\mathbf{A}_2$ es combinación lineal de los vectores anteriores a él y por tanto el conjunto original es linealmente dependiente. Debemos descartar al vector $\mathbf{A}_2$. Realizando solo el cálculo:

$$\mathbf{A}_3 \bullet \mathbf{B}_1 = 1;$$

Obtenemos:

$$x_{13} = 1/2$$

Y así:

$$\mathbf{B}_3 = \mathbf{A}_3 - x_{13}\mathbf{B}_1 = \begin{bmatrix} 1/2 & 1 \\ -1/2 & 2 \end{bmatrix}_\diamond$$

Ejercicio 21

Aplique el algoritmo anterior al conjunto formado por las matrices:

$$\mathbf{A}_1 = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \mathbf{A}_2 = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}, \mathbf{A}_3 = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$$

Sugerencia
Siga el algoritmo; no hay de otra.

8.12. Existencia de Bases Ortonormales

EL resultado anterior se conoce como el procedimiento de Gram-Schmidt, y permite cambiar una base por una base ortogonal, de donde es fácil obtener una base ortonormal:

Teorema 8.8

Todo espacio lineal de matrices $m \times n$ posee una base ortonormal

Demostración
Sabemos que todo espacio lineal posee una base, digamos $\{\mathbf{A}_1, \ldots, \mathbf{A}_k\}$. Sean lo $\mathbf{B}_j$ los las matrices obtenidas de aplicar el proceso de Gram-Schmidt. Los espacios generado por los $\mathbf{B}_i$ y los $\mathbf{A}_i$ son iguales pues las matrices $\mathbf{B}_i$ son combinaciones lineales de los vectores $\mathbf{A}_i$ y viceversa. Así, los conjuntos generados son los mismos. Si las matrices $\mathbf{A}_i$ forman un conjunto linealmente independiente, también lo debe ser el conjunto formado por los $\mathbf{B}_i$. Por tanto, el conjunto formado por los $\mathbf{B}_i$ es también una base para el espacio generado por los $\mathbf{A}_i$.$\diamond$

Una ventaja inmediata de una base ortonormal $\{\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_k\}$ es que si $\mathbf{A}$ es una combinación lineal de las matrices anteriores entonces:

$$\mathbf{A} = (\mathbf{A} \bullet \mathbf{A}_1)\,\mathbf{A}_1 + (\mathbf{A} \bullet \mathbf{A}_2)\,\mathbf{A}_2 + \cdots + (\mathbf{A} \bullet \mathbf{A}_k)\,\mathbf{A}_k$$

es decir, que no hace falta utilizar el proceso de eliminación gaussiana para determinar los coeficientes, si no simplemente hacer un producto punto con el vector correspondiente.
Una consecuencia también inmediata es que se puede completar un conjunto linealmente independiente y ortogonal a una base ortogonal en un espacio lineal de matrices $m \times n$. Una forma conviente es aplicar el proceso de Gram-Schmidt al conjunto li y ortogonal aumentado con una base cualquiera del espacio y detener el proceso hasta obtener el primer vector cero. El resultado esperado debe ser el conjunto inicial adicionado con algunos vectores. Este conjunto debe ser una base para el espacio completo.

8.13. Proyección ortogonal

Teorema 8.9

Sea $\mathbf{Y}$ una matriz $m \times n$ y un espacio lineal V de dimensión r, ambos dentro de un espacio lineal U. Entonces, existe una **única** matriz $\mathbf{Z}$ en V tal que $(\mathbf{Y} - \mathbf{Z}) \perp V$. Si $r = 0$ entonces $\mathbf{Z} = 0$, y si $r > 0$ entonces $\mathbf{Z}$ se puede expresar como

$$\mathbf{Z} = c_1\mathbf{X}_1 + \ldots + c_r\mathbf{X}_r,$$

donde $\{\mathbf{X}_1, \ldots, \mathbf{X}_r\}$ forman una base ortonormal de V y $c_i = \mathbf{Y} \bullet \mathbf{X}_i$ para $i = 1, \ldots, r$. Además, $\mathbf{Z} = \mathbf{Y}$ si y sólo si $\mathbf{Y} \in V$. La matriz $\mathbf{Z}$ se llamará *la proyección ortogonal* de $\mathbf{Y}$ sobre V y cumple que $d(Z, Y) \leq d(X, Y)$ para toda $\mathbf{X}$ en V y hay igualdad si y sólo si $\mathbf{Z} = \mathbf{X}$.

Demostración

Si $r = 0$ entonces $\dim(V) = 0$, y por tanto $V = \{\mathbf{0}\}$. Para $\mathbf{Z} = \mathbf{0}$ se tiene $(\mathbf{Y} - \mathbf{Z}) \perp V$. Y es claramente la única matriz en V que cumple esto.

Si $r > 0$ sea $\{\mathbf{X}_1, \ldots, \mathbf{X}_r\}$ una base ortonormal de V y definamos $c_i = \mathbf{Y} \bullet \mathbf{X}_i$ para $i = 1, \ldots, r$ y $\mathbf{Z} = \sum_{i=1}^{r} c_i\mathbf{X}_i$. Claramente, $\mathbf{Z} \in V$ y

$$\left(\mathbf{Y} - \sum_{i=1}^{r} c_i\mathbf{X}_i\right) \bullet \mathbf{X}_j = \mathbf{Y} \bullet \mathbf{X}_j - c_j = c_j - c_j = 0$$

para cada $j = 1, \ldots, r$. Y por tanto, $(\mathbf{Y} - \mathbf{Z}) \perp V$. Si ahora $\mathbf{X} \in V$ y cumple $(\mathbf{Y} - \mathbf{X}) \perp V$ entonces:

$$\begin{aligned} (\mathbf{X} - \mathbf{Z}) \bullet (\mathbf{X} - \mathbf{Z}) &= (\mathbf{X} - \mathbf{Y} + \mathbf{Y} - \mathbf{Z}) \bullet (\mathbf{X} - \mathbf{Z}) \\ &= -(\mathbf{Y} - \mathbf{X}) \bullet (\mathbf{X} - \mathbf{Z}) + (\mathbf{Y} - \mathbf{Z}) \bullet (\mathbf{X} - \mathbf{Z}) \\ &= -0 + 0 = 0 \end{aligned}$$

Esto debido a que $\mathbf{X} - \mathbf{Z} \in V$ y a que $(\mathbf{Y} - \mathbf{Z}) \bullet (\mathbf{X} - \mathbf{Z}) = 0$. Por tanto, $\mathbf{X} - \mathbf{Z} = \mathbf{0}$ y de allí que $\mathbf{X} = \mathbf{Z}$, haciendo que $\mathbf{Z}$ sea el único vector en V que cumple $(\mathbf{Y} - \mathbf{Z}) \perp V$. Si $\mathbf{X}$ es una matriz en V entonces $\mathbf{X} - \mathbf{Z}$ también está en V y por tanto $\mathbf{X} - \mathbf{Z}$ será ortogonal a $\mathbf{Y} - \mathbf{Z}$. De donde deducimos que

$$\|\mathbf{X} - \mathbf{Y}\|^2 = \|\mathbf{X} - \mathbf{Z}\|^2 + \|\mathbf{Y} - \mathbf{Z}\|^2$$

de donde

$$d(\mathbf{X}, \mathbf{Y})^2 = d(\mathbf{X}, \mathbf{Z})^2 + d(\mathbf{Y}, \mathbf{Z})^2$$

y por tanto $d(\mathbf{Z}, \mathbf{Y}) \leq d(\mathbf{X}, \mathbf{Y})$. Y esto se cumple con igualdad si y sólo si $\mathbf{X} = \mathbf{Z}$ ■

Ejercicio 22

Considere el espacio lineal formado por todas las soluciones al sistema homogéneo:

$$\begin{aligned} x + y + z - w &= 0 \\ x - y - z + w &= 0 \end{aligned}$$

y el vector $\mathbf{d} =< 1, 3, 2, 1 >$.

- Usando el orden primero x, luego y, luego z, y por último w, encuentre una base para el espacio solución.
- Ortogonolice la base encontrada.

- Usando la base encontrada, determine la proyección ortogonal de $\mathbf{d}$ sobre tal espacio.
- Usando el orden primero y, luego w, luego y, y por último z, encuentre una base para el espacio solución.
- Ortogonolice la base nueva base.
- Usando la nueva base encontrada, determine la proyección ortogonal de $\mathbf{d}$ sobre tal espacio.

8.14. Descomposición QR

Sea $\mathbf{A}$ una matriz $m \times k$ de rango columna completo. Esto es, una matriz de rango k. Simbolicemos por $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k$ las columnas de la matriz $\mathbf{A}$. Por el resultado anterior, deben existir escalares $x_{i,j}(i < j = 1, \ldots, k)$ tales que los vectores columna $\mathbf{b}_1, \ldots, \mathbf{b}_k$ definidos en forma recursiva por las igualdades:

$$\begin{aligned}
\mathbf{b}_1 &= \mathbf{a}_1 \\
\mathbf{b}_2 &= \mathbf{a}_2 - x_{1,2}\mathbf{b}_1 \\
&\vdots \\
\mathbf{b}_j &= \mathbf{a}_j - x_{1,j}\mathbf{b}_1 - \cdots - x_{j-1,j}\mathbf{b}_{j-1} \\
&\vdots \\
\mathbf{b}_k &= \mathbf{a}_k - x_{1,k}\mathbf{b}_1 - \cdots - x_{k-1,k}\mathbf{b}_{k-1}
\end{aligned}$$

o por las igualdades:

$$\begin{aligned}
\mathbf{a}_1 &= \mathbf{b}_1 \\
\mathbf{a}_2 &= \mathbf{b}_2 + x_{1,2}\mathbf{b}_1 \\
&\vdots \\
\mathbf{a}_j &= \mathbf{b}_j + x_{1,j}\mathbf{b}_1 + \cdots + x_{j-1,j}\mathbf{b}_{j-1} \\
&\vdots \\
\mathbf{a}_k &= \mathbf{b}_k + x_{1,k}\mathbf{b}_1 + \cdots + x_{k-1,k}\mathbf{b}_{k-1}
\end{aligned}$$

forman un conjunto ortogonal. Sea $\mathbf{B}$ la matriz cuya columna i es el vector $\mathbf{b}_i$, y sea $\mathbf{X}$ la matriz $k \times k$ triangular superior unitaria cuyo elemento (i, j) es $x_{i,j}$ entonces:

$$\mathbf{A} = \mathbf{B}\,\mathbf{X}$$

Si se define la matriz $\mathbf{D}$ como

$$\begin{aligned}
\mathbf{D} &= \operatorname{diag}\left(\|\mathbf{b}_1\|^{-1}, \|\mathbf{b}_2\|^{-1}, \ldots, \|\mathbf{b}_k\|^{-1}\right), \\
\mathbf{E} &= \operatorname{diag}\left(\|\mathbf{b}_1\|, \|\mathbf{b}_2\|, \ldots, \|\mathbf{b}_k\|\right)
\end{aligned}$$

y como

$$\mathbf{Q} = \mathbf{B}\,\mathbf{D} \text{ y } \mathbf{R} = \mathbf{E}\,\mathbf{X}$$

entonces

$$\mathbf{A} = \mathbf{Q}\mathbf{R}$$

es una factorización de $\mathbf{A}$ en dos matrices: la primera formada por columnas que son ortonormales y la segunda triangular superior con elementos de la diagonal principal positivos. Esta factorización se conoce como la *factorización QR* de $\mathbf{A}$ y es muy importante en métodos numéricos.

Ejemplo 7

Determine la factorización $\mathbf{QR}$ de la matriz:

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 1 & 5 \\ 1 & 6 \end{bmatrix}$$

Solución: Trabajemos con las columnas de $\mathbf{A}$:

$$\mathbf{a}_1 = (1,1,1,1,1,1)',\ \mathbf{a}_2 = (1,2,3,4,5,6)'$$

Así:

$$\mathbf{b}_1 = \mathbf{a}_1$$

Calculando:

$$\mathbf{b}_1 \bullet \mathbf{b}_1 = 6 \text{ y } \mathbf{a}_2 \bullet \mathbf{b}_1 = 21$$

Obtenemos:

$$x_{12} = 7/2$$

Y así:

$$\mathbf{b}_2 = \mathbf{a}_2 - x_{12}\,\mathbf{b}_1 = (-5/2, -3/2, -1/2, 1/2, 3/2, 5/2)'$$

Calculando:

$$\mathbf{b}_2 \bullet \mathbf{b}_2 = 35/2$$

Por tanto

$$\mathbf{Q} = [\mathbf{b}_1, \mathbf{b}_2]\mathbf{D} = \begin{bmatrix} 1 & -5/2 \\ 1 & -3/2 \\ 1 & -1/2 \\ 1 & 1/2 \\ 1 & 3/2 \\ 1 & 5/2 \end{bmatrix} \begin{bmatrix} \sqrt{6}/6 & 0 \\ 0 & \sqrt{70}/35 \end{bmatrix}$$

$$\mathbf{Q} = \begin{bmatrix} 1/6\sqrt{6} & -1/14\sqrt{70} \\ 1/6\sqrt{6} & -3/70\sqrt{70} \\ 1/6\sqrt{6} & -1/70\sqrt{70} \\ 1/6\sqrt{6} & 1/70\sqrt{70} \\ 1/6\sqrt{6} & 3/70\sqrt{70} \\ 1/6\sqrt{6} & 1/14\sqrt{70} \end{bmatrix}$$

$$\mathbf{R} = \begin{bmatrix} \sqrt{6} & 0 \\ 0 & 1/2\sqrt{70} \end{bmatrix} \begin{bmatrix} 1 & x_{12} \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \sqrt{6} & \frac{7}{2}\sqrt{6} \\ 0 & \frac{1}{2}\sqrt{70} \end{bmatrix}$$

Ejercicio 23

Determine la factorización $\mathbf{QR}$ de la matriz:

$$\mathbf{A} = \begin{bmatrix} 2 & 1 \\ -1 & 1 \\ 2 & 2 \\ -1 & 0 \end{bmatrix}$$

8.15. Aplicación de la factorización QR

Teorema 8.10

Si $\mathbf{A}$ es una matriz cuyas columnas son linealmente independientes y si $\mathbf{A} = \mathbf{QR}$ es una factorización $\mathbf{QR}$ de $\mathbf{A}$. Entonces la única solución $\tilde{\mathbf{x}}$ de $\mathbf{Ax} = \mathbf{b}$ por mínimos cuadrados se expresa teóricamente como

$$\tilde{\mathbf{x}} = \mathbf{R}^{-1}\mathbf{Q}^T\mathbf{b}$$

Ejemplo 8
Aplique la factorización QR para resolver por mínimos cuadrados el sistema:

$$\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 1 & 5 \\ 1 & 6 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 1 \\ 2 \\ 0 \\ 1 \\ 1 \\ 2 \end{bmatrix}$$

Solución Por el ejemplo anterior, la factorización de la matriz de coeficientes es:

$$\mathbf{A} = \mathbf{QR} = \begin{bmatrix} 1/6\sqrt{6} & -1/14\sqrt{70} \\ 1/6\sqrt{6} & -3/70\sqrt{70} \\ 1/6\sqrt{6} & -1/70\sqrt{70} \\ 1/6\sqrt{6} & 1/70\sqrt{70} \\ 1/6\sqrt{6} & 3/70\sqrt{70} \\ 1/6\sqrt{6} & 1/14\sqrt{70} \end{bmatrix} \begin{bmatrix} \sqrt{6} & \frac{7}{2}\sqrt{6} \\ 0 & \frac{1}{2}\sqrt{70} \end{bmatrix}$$

Siguiendo la estrategia descrita en el último resultado la solución por el método de mínimos cuadrados se calcula por la fórmula:

$$\tilde{x} = \mathbf{R}^{-1}\mathbf{Q}^T\mathbf{b}$$

Así

$$\mathbf{Q}^T = \begin{bmatrix} 1/6\sqrt{6} & 1/6\sqrt{6} & 1/6\sqrt{6} & 1/6\sqrt{6} & 1/6\sqrt{6} & 1/6\sqrt{6} \\ -1/14\sqrt{70} & -3/70\sqrt{70} & -1/70\sqrt{70} & 1/70\sqrt{70} & 3/70\sqrt{70} & 1/14\sqrt{70} \end{bmatrix}$$

y

$$\mathbf{R}^{-1} = \begin{bmatrix} 1/6\sqrt{6} & -1/10\sqrt{70} \\ 0 & 1/35\sqrt{70} \end{bmatrix}$$

y por tanto,

$$\tilde{x} = \mathbf{R}^{-1}\mathbf{Q}^T\mathbf{b} = \begin{bmatrix} \frac{13}{15} \\ \frac{3}{35} \end{bmatrix} \diamond$$

Ejercicio 24

Aplique la factorización QR para resolver por mínimos cuadrados el sistema:

$$\begin{bmatrix} 2 & 1 \\ -1 & 1 \\ 2 & 2 \\ -1 & 0 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 1 \\ 2 \\ 0 \\ 1 \end{bmatrix}$$

8.16. Uso de la TI

La calculadora TI 89 o TI Voyage 200 puede ser utilizada para la factorización **QR**. La ventaja de la rutina numérica programada es que da información sobre el espacio nulo de la matriz de entrada cuando las columnas de la matriz no son linealmente independientes. Considere por ejemplo la matriz

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$$

Las figuras 3 y 4 muestran la entrada y los cálculos en la TI. El punto a señalar es que habiendo ceros en la diagonal de **R** se indica que los vectores correspondientes en **Q** están en el espacio nulo de la matriz **A**. De

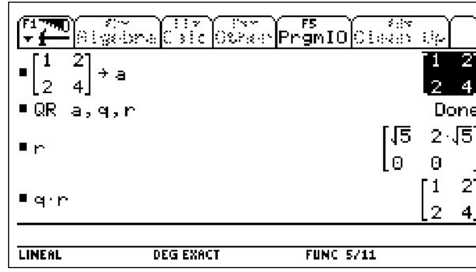


Figura 3: Descomposición $\mathbf{QR}$ de $\mathbf{A}$ en la TI.

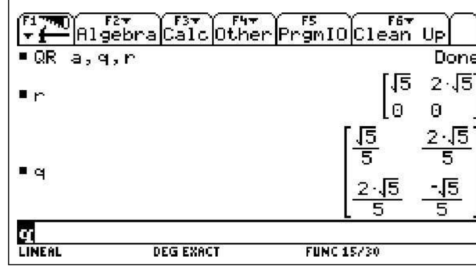


Figura 4: La naturaleza de los ceros en la diagonal de $\mathbf{R}$.

hecho, todos los vectores correspondientes a ceros en la diagonal de $\mathbf{R}$ forman la base para el espacio nulo de $\mathbf{A}$.

Ejemplo 9

Determine la distancia de $P(0,-2,-1)$ al espacio que generan los vectores:

$$\left\{ \mathbf{v}_1 = \begin{pmatrix} -2 \\ 3 \\ -1 \end{pmatrix}, \mathbf{v}_2 = \begin{pmatrix} 1 \\ 5 \\ -3 \end{pmatrix} \right\}$$

Con los vectores $\mathbf{v}_1$ y $\mathbf{v}_2$ formamos la matriz $\mathbf{A}$ a la cual le aplicamos la factorización QR:

$$\mathbf{A} = [\mathbf{v}_1\,\mathbf{v}_2] = \mathbf{Q}\,\mathbf{R} = \begin{bmatrix} -1/7\sqrt{14} & 17/1365\sqrt{1365} \\ 3/14\sqrt{14} & 4/273\sqrt{1365} \\ 1/14\sqrt{14} & -2/105\sqrt{1365} \end{bmatrix} \cdot \begin{bmatrix} \sqrt{14} & 5/7\sqrt{14} \\ 0 & 1/7\sqrt{1365} \end{bmatrix}$$

De donde los coeficientes de Fourier del vector que representa a P,$\mathbf{v} =< 0,-2,1 >$ son:

$$[\mathbf{v}]_{\mathbf{Q}} = \mathbf{Q}^{\mathrm{T}} \cdot \mathbf{v} = \begin{pmatrix} -\frac{1}{2}\sqrt{14} \\ -\frac{2}{195}\sqrt{1365} \end{pmatrix}$$

y de allí que la proyección de $\mathbf{v}$ a $C(\mathbf{A})$ es:

$$\mathbf{v}_{C(\mathbf{A})} = \mathbf{Q} \cdot [\mathbf{v}]_{\mathbf{Q}} = \begin{pmatrix} \frac{161}{195} \\ -\frac{133}{78} \\ -\frac{7}{30} \end{pmatrix}$$

Por tanto, la distancia de $\mathbf{v}$ a $C(\mathbf{A})$ es

$$d(\mathbf{v}, C(\mathbf{A})) = \left|\left|\mathbf{v} - \mathbf{v}_{C(\mathbf{A})}\right|\right| = \left|\left| \begin{pmatrix} -\frac{161}{195} \\ -\frac{23}{78} \\ -\frac{23}{30} \end{pmatrix} \right|\right| = \frac{23}{390}\sqrt{390}$$

Ejemplo 10
Determine la distancia de $P(0,-2,-1,1)$ al conjunto de soluciones del sistema

$$\mathbf{B}\,\mathbf{x} = \begin{bmatrix} 2 & 1 & 3 & 1 \\ -1 & 0 & 2 & 1 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Este problema se resuelve como el anterior, pero debemos encontrar un conjunto generador. Para ello tenemos dos alternativas: o seguir el proceso descrito en lecturas anteriores o calcular lo que se conoce como el espacio nulo de $\mathbf{B}$: Aquellos vectores $\mathbf{x}$ tal que $\mathbf{Bx} = \mathbf{0}$. En general, los conjuntos generadores obtenidos por ambos procesos son diferentes pero ambos son base para el mismo conjunto. Si utilizamos la alternativa del espacio nulo tenemos:

$$\text{nullspace}(\mathbf{B}) = \text{Gen}\left\{ \begin{pmatrix} 2 \\ -7 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -3 \\ 0 \\ 1 \end{pmatrix} \right\}$$

Así el problema se ha transformado en encontrar la distancia de $P(=\mathbf{v})$ a $C(\mathbf{A})$ donde $\mathbf{A}$ es la matriz cuyas columnas son el generador de nullspace($\mathbf{B}$). Se deja como ejercicio comprobar que:

$$\mathbf{A} = \begin{bmatrix} 2 & 1 \\ -7 & -3 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathbf{QR} = \begin{bmatrix} 1/9\,\sqrt{6} & 4/585\,\sqrt{390} \\ -7/18\,\sqrt{6} & -1/1170\,\sqrt{390} \\ 1/18\,\sqrt{6} & -23/1170\,\sqrt{390} \\ 0 & 3/65\,\sqrt{390} \end{bmatrix} \cdot \begin{bmatrix} 3\,\sqrt{6} & 23/18\,\sqrt{6} \\ 0 & 1/18\,\sqrt{390} \end{bmatrix}$$

$$\mathbf{v}_{C(\mathbf{A})} = \begin{pmatrix} 43/65 \\ -19/65 \\ -47/65 \\ -14/65 \end{pmatrix}$$

$$d(\mathbf{v}, C(\mathbf{A})) = 1/65\,\sqrt{4615}$$

Ejemplo 11
Determine la matriz $\mathbf{X}$ que mejor resuelve el sistema $\mathbf{A}\,\mathbf{X} = \mathbf{B}$ para

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 0 & -1 \\ 1 & 1 \end{bmatrix}, \ \mathbf{B} = \begin{bmatrix} 1 & 2 & -1 \\ 0 & -1 & 8 \\ 1 & 1 & 2 \end{bmatrix}$$

La factorización QR de $\mathbf{A}$ es

$$\mathbf{A} = \mathbf{QR} = \begin{bmatrix} 1/2\,\sqrt{2} & 1/6\,\sqrt{6} \\ 0 & -1/3\,\sqrt{6} \\ 1/2\,\sqrt{2} & -1/6\,\sqrt{6} \end{bmatrix} \begin{bmatrix} \sqrt{2} & 3/2\,\sqrt{2} \\ 0 & 1/2\,\sqrt{6} \end{bmatrix}$$

Por tanto, la matriz $\mathbf{X}$ que minimiza el error cuadrático es:

$$\mathbf{X} = \mathbf{R}^{-1}\mathbf{Q}'\mathbf{B} = \begin{bmatrix} 1 & 0 & 10 \\ 0 & 1 & -19/3 \end{bmatrix}$$

Y el error cometido con $\mathbf{X}$ es 25/3 ■